Accepted Manuscript

Lossless Coding of Light Field Images based on Minimum-Rate Predictors

João M. Santos, Pedro A.A. Assuncao, Luis A. da Silva Cruz, Luis M.N. Tavora, Rui Fonseca-Pinto, Sergio M.M. Faria

PII: S1047-3203(18)30051-8
DOI: https://doi.org/10.1016/j.jvcir.2018.03.003
Reference: YJVCI 2150

To appear in: *J. Vis. Commun. Image R.*

Received Date: 31 August 2017
Accepted Date: 3 March 2018

Please cite this article as: J.M. Santos, P.A.A. Assuncao, L.A. da Silva Cruz, L.M.N. Tavora, R. Fonseca-Pinto, S.M.M. Faria, Lossless Coding of Light Field Images based on Minimum-Rate Predictors, *J. Vis. Commun. Image R.* (2018), doi: https://doi.org/10.1016/j.jvcir.2018.03.003

Lossless Coding of Light Field Images based on Minimum-Rate Predictors

João M. Santos[a,b,*], Pedro A. A. Assuncao[a,c], Luis A. da Silva Cruz[a,b], Luis M. N. Tavora[c], Rui Fonseca-Pinto[a,c], Sergio M. M. Faria[a,c]

[a] *Instituto de Telecomunicações, Portugal*
[b] *Department of Electrical and Computer Engineering, Faculty of Sciences and Technology, University of Coimbra, Coimbra 3030-290, Portugal*
[c] *Escola Superior de Tecnologia e Gestão (ESTG), Instituto Politécnico de Leiria, Leiria 2411-901, Portugal*

Abstract

Recent developments in light field acquisition and computational photography are driving new research efforts on light field encoding methods, capable of exploiting the specific features of this type of visual data. This paper presents a research study of lossless light field image compression, using Minimum-Rate Predictors (MRP) and mainstream image and video encoders. The research is focused on three light field representation formats: lenslet images, stack of sub-aperture images and epipolar images. The main contributions of this work are the 'Spiral-blackend' serialization method and the use of MRP for the lossless compression of light fields with joint encoding of RGB data. The results show that the lenslet format yields lower compression efficiencies than other formats. Furthermore, it is demonstrated that the MRP algorithm consistently outperforms HEVC-RExt, JPEG2000, JPEG-LS and CALIC when light fields are represented by either a stack of sub-aperture or epipolar images.

Keywords: Image coding, lossless compression, light field coding.

1. Introduction

Light field (LF) imaging, also known as plenoptic, holoscopic or integral imaging, a concept first introduced by Lippman in 1908 [1], has started to receive more attention once consumer-grade devices, such as those from Lytro (https://www.lytro.com/) and Raytrix (https://www.raytrix.de/) appeared in the market. These devices are capable of recording information about both the intensity and direction of the light rays reaching the camera, through the use of a specially designed optical system placed right in front of the image sensor. Other LF acquisition set-ups based on arrays of cameras, and therefore with different sizes, geometries and imaging characteristics, have also been developed mainly for research purposes, such as the camera array used to capture part of the LF data available at The (New) Stanford Light Field Archive [2].

LF cameras capture significantly more visual information from scenes than regular cameras, since besides light intensity from each point in the scene, they also record information related to the direction of the light rays travelling towards the sensor. One type of LF camera is built around an optical system that comprises a main lens and a micro-lenses array (MLA), placed in the optical path between the main lens and the light sensor [3], as illustrated in Figure 1. In such arrangement, each micro-lens of the MLA captures the light intensity from each point in the scene, coming from different perspectives. Accordingly, as represented in Figure 1, each sample acquired by the image sensor corresponds to the intensity of a light-ray coming from a point in the visual scene with a certain direction. This results in a 4D representation of the sampled LF in the form of a matrix of micro-images, each one corresponding to the light samples captured through a single micro-lens, where the spatial position of each sample implicitly encodes directional information. After adequate calibration and rectification, different views of the visual scene can be obtained from the 4D LF by extracting pixels located at the same spatial position in each micro-image. The whole set of views may comprise a stack of sub-aperture images, providing another representation format for the LF. A more detailed description and analysis of LF acquisition systems can be found in [4, 5, 6].

LF imaging is a primary enabler of computational photography, since the captured raw visual data can be processed *a posteriori*, in order to extract other image related data and render images with different characteristics. Amongst the operations that can be performed on the raw data, one can quote extraction of depth maps [7] and rendering of images with different focal planes, depth-of-field or viewing perspectives [3, 8, 9]. These technical possibilities make LF images useful in many fields of application, from medical imaging to the volumetric recording of prized sculptures and other cultural heritage artefacts [10]. However, this flexibility and the new functionalities come with a cost: the need to process and store the very large amount of

*Corresponding author
E-mail addresses: joao.santos@co.it.pt (João M. Santos)., amado@co.it.pt (Pedro A. A. Assuncao)., luis.cruz@co.it.pt (Luis A. da Silva Cruz)., luis.tavora@ipleiria.pt (Luis M. N. Tavora)., rui.pinto@ipleiria.pt (Rui Fonseca-Pinto)., sergio.faria@co.it.pt (Sergio M. M. Faria).